Erickson ＆ Kawahara (2016) Articulatory correlates of metrical structure: Studying jaw displacement patterns *Linguistic Vanguard* 2.

Addendum to the ROA version:

Although this paper does not directly address issues in Optimality Theory, we find patterns that are very similar to what has been proposed in the OT literature. For example, we find that English metrical prominence correlates with degrees of jaw displacement. This mapping reminds us of Harmonic Alignment from Prince and Smolensky's work. In French, we find large jaw opening at right phrase edges, which resembles EndRule-R. In Japanese, we find large jaw opening at both edges of a phrase, which means that EndRule-L and EndRule-R are operative in Japanese. We also show that Japanese pitch accent does not result in large jaw opening, which shows that the mapping from metrical prominence to large jaw opening is governed by a <u>violable</u> principle. Overall, generally, we have been pleasantly surprised to find that the phonology-phonetics mapping principles we find in our jaw movement data resemble those principles proposed in the OT literature.

Articulatory correlates of metrical structure:
Studying jaw displacement patterns

Abstract

Previous phonetic studies of metrical prominence have primarily focused on its acoustic manifestations, including pitch, intensity, duration, spectral tilt, etc. In this paper we outline our new research program in which we explore jaw displacement patterns as another articulatory reflex of metrical prominence. We present our studies of English and Japanese in some detail, which show that jaw movement patterns are neither flat nor random, but instead the degrees of jaw displacement correlate well with metrical prominence. Based on these results, we argue that there are at least two articulators to express metrical prominence: the larynx and the jaw. Our aim is not so much to object to looking at the acoustic manifestations of metrical structures or other articulation-based approaches; we instead would like to encourage other researchers to investigate metrical structure in terms of jaw movement as well.

1. Introduction

The theory of Metrical Phonology (Liberman & Prince 1977; Selkirk 1980a,b *et seq*.) posits that languages organize their utterances into hierarchical layers of metrical units: e.g., mora, syllable, foot, Prosodic Word, Prosodic Phrase and Utterance. Phonologically speaking, metrical structures determine stress location, tonal alignment, segmental phonotactics, and domains of rule applications. Previous phonetic research has shown that these metrical constituents also affect phonetic implementation patterns. For example, previous studies of Japanese intonation show that prosodic phrasing determines the distribution of tonal rises, the location of pitch reset (i.e. the domain of downstep), and the domain of general declination (e.g. Kawahara & Shinya 2008; Pierrehumbert & Beckman 1988; Poser 1984). Phonetic work has also revealed many acoustic cues that are associated with stress in many languages: pitch, intensity, duration, spectral tilt and others (Beckman 1986; Fry 1955; Plag et al. 2011 and many others; see also Eriksson 2009).

Thanks to this research tradition, we now have a fairly good understanding of how languages can and cannot differ in terms of prosodic organization (see especially Jun 2005, 2014), and how these metrical structures are realized acoustically across different languages. One starting point of our research, however, was the observation that much work on the phonetic realization of metrical structures has focused on those properties that are controlled

by the larynx.[1]

In this paper we outline our new research program to investigate another dimension in which metrical structure manifests itself, in particular jaw displacements. The essence of our proposal is that there are at least two "prosodic articulators": the larynx, as we already know, and also the jaw. The fact that the jaw is relevant to prosodic manifestation can be illustrated with a very simple example (pointed out to us by Doug Whalen). In English, unstressed syllables—or metrically reduced syllables—are realized as schwas, which involve very small jaw displacement. There is a clear sense in which prosody manifests itself in terms of jaw displacement in this example.

Beyond this simple example, we hope to show that jaw articulation is a reliable—or at the very least interesting—measure of metrical prominence and rhythm. Our illustration of this research enterprise will unfold as follows. In section 2, we first illustrate our hypothesis by reviewing earlier studies which show that jaw displacement increases with contrastive emphasis. In section 3, we demonstrate how jaw displacement patterns reflect sentential metrical prominence in English. Section 4 discusses jaw movement patterns in Japanese, which is often considered to lack stress. We demonstrate that metrically prominent syllables nevertheless show large jaw opening. In section 5, we address some alternative analyses of what is discussed in sections 3 and 4. In section 6, we discuss our preliminary results from other languages. Section 7 discusses further implications of our study, including L1 transfer of jaw displacement patterns to L2 acquisition. As per the spirit of this journal, this paper should be taken as a declaration of a new, admittedly tentative, research program rather than a fully developed defense of a completely fleshed-out theory. The paper raises more questions than it answers, but we believe that it opens up many research opportunities for future studies.

2. Contrastive emphasis = metrically prominent = large jaw opening

We start illustrating our hypothesis—jaw displacement is one way in which metrical prominence manifests itself—with an old and well-known observation. Contrastive emphasis on a word makes a particular syllable more metrically prominent (Féry 2013; Ladd 2008). Such syllables with contrastive emphasis are usually pronounced with higher pitch, longer duration, and larger intensity (e.g. Eady et al. 1986; Selkirk & Katz 2011).

A large number of articulatory studies of English have reported increased jaw

---

[1] This is not to suggest that no one has investigated the articulatory aspects of metrical structure. Cho (2006) for example explored the effects of prosodic structure on lip kinematics. Tilsen (2009) investigated the relationship between variability in articulatory gestures and variability in higher rhythmic units, including syllables and feet. Work on articulatory strengthening at domain-initial positions (e.g. Cho et al. 2007) and domain-final lengthening (e.g. Edwards et al. 1991) is also relevant in that it has investigated articulatory correlates of phrasal patterns at initial and final positions (for which see section 5 for more discussion). The π-gesture model (Byrd & Saltzman 2003; Byrd et al. 2006) explores the mechanism behind the articulation of phrase-edge lengthening. Our research program is obviously inspired by this body of work, but differs from it by specifically examining jaw displacement patterns as general articulatory correlates of metrical prominence. We hasten to add that we do not intend to argue against these research programs.

displacement for contrastively emphasized syllables.[2] Increased jaw displacement has also been reported for emphasized words in other languages, such as Japanese (Erickson et al. 2000; Maekawa et al. 1998) and French (Loevenbruck 1999).

These studies have been based on data from x-ray microbeam or EMA (ElectroMagnetic Articulography). Figure 1 shows the place of measurement of jaw displacement recorded using EMA (the bottom panel). In this method, jaw displacement measurements are usually measured from the bite plane (occlusal plane) to the maximum point of jaw opening during each syllable.[3]
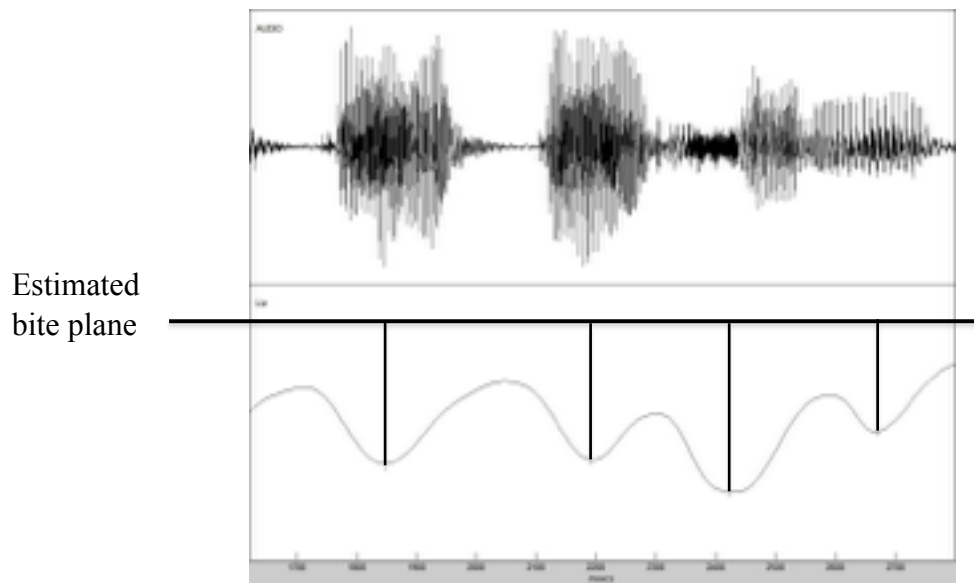

Figure 1: Waveform (top) and jaw displacement measurement using EMA (bottom).

As a representative case study, Erickson (2002) studied x-ray microbeam recordings of emphasized words and unemphasized words containing /aɪ/ and other vowels. Erickson (2003a) likewise studied 44 speakers from the x-ray microbeam database, targeting words containing /æ/ and /i/. The results from Erickson (2002), partially reproduced in Figure 2, show that jaw opening, measured at its peak displacement, is significantly lower for

---

[2] They include Beckman & Edwards (1994), de Jong (1995), Erickson (1998, 2002, 2003a,b), Erickson & Honda (1996), Erickson & Fujimura (1992, 1996), Erickson et al. (1994, 1999), Harrington et al. (2000), Kent & Netsell (1971), Macchi (1985, 1988), Menezes (2003, 2004, 2015), Oshimat & Gracco (1992), Perkell (1969), Stone (1981), Summers (1987), Westbury & Fujimura (1989), and others.

[3] The work summarized in this paper focuses on the vertical displacement of the jaw, without considering its rotational movement. Whether metrical prominence correlates with the amount of vertical displacement or the trajectory of rotational movement is an interesting question, but we leave this issue for future studies.

emphasized /aɪ/ than for unemphasized /aɪ/.[4] Table 1, taken from Erickson (2003a), shows average differences between emphasized and unemphasized /æ/ and /i/ in English, showing that for both vowels, emphasized vowels show larger jaw displacement.
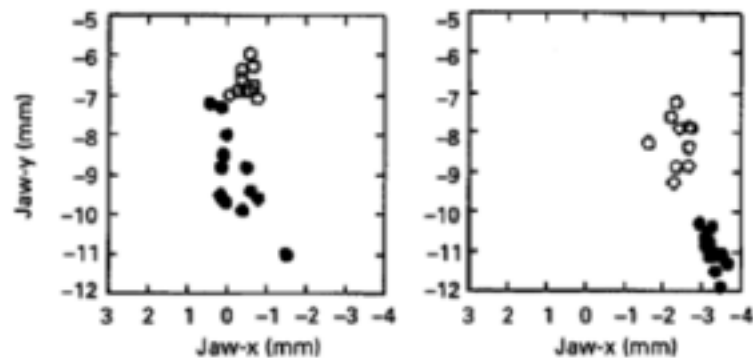

Figure 2: Data for two speakers on emphasized and non-emphasized /aɪ/. Filled circles indicate emphasized vowels; unfilled circles represent non-emphasized vowels. The lower the y-axis value, the larger jaw displacement. Reproduced from Erickson (2002) with permission.

Table 1: Average differences in jaw displacement values between emphasized and unemphaized vowels, broken down by vowel quality and speaker gender. All the differences are significant at *p*<.01 level. Taken from Erickon (2003a).

Vowel type	speaker gender	N	Differences
/i/	m	16	-1.6
/i/	f	26	-1.2
/æ/	m	18	-4.1
/æ/	f	26	-3.5

We can interpret these results as follows: those elements that receive contrastive focus are metrically strong phonologically (Féry 2013; Ladd 2008), and hence they manifest themselves with large jaw opening.

One prediction that the present theory makes is that everything else being equal, those elements with contrastive focus should show larger jaw opening, given the fairly uncontroversial assumption that contrastive focus assigns metrical prominence. This prediction is supported by Menezes et al. (2003), in which the location of contrastive focus is systematically varied within a phrase like "five-nine-five Pine Street". They observe that those syllables that are contrastively emphasized within this phrase show large jaw opening. As the metrical structure of a sentence changes due to contrastive emphasis, so does the jaw movement pattern, the governing principle being that metrical prominence is expressed via

---

[4] Depending on the vowel height, the jaw opens more or less (e.g. Keating et al 1994; see section 3.2); for low vowels, the jaw opens more than for high vowels. However, regardless of the vowel height, the mouth opens more for emphasized words (Erickson 1998, 2002). In addition, the tongue dorsum moves more in the phonological direction of the vowel. Acoustically, emphasized low vowels show higher F1 and lower F2, whereas emphasized high vowels show lower F1 and higher F2 (Erickson 2002).

large jaw opening.

3. Jaw displacement patterns and sentential metrical prominence in English
3.1. Metrical prominence and jaw opening

One may quibble from the previous discussion that jaw movement becomes important only when contrastive focus comes in. However, some studies show that the role of jaw movement pattern is more pervasive: jaw displacement patterns may in fact reflect sentential metrical prominence in English sentences in general.

To model sentential prominence patterns, Metrical Phonology posits that syllables are hierarchically organized, and within each level, some syllables are designated as stronger than the others. This prominence is formally expressed by adding a grid mark for that level. The result is that each syllable is assigned a different number of grid marks, which represents relative prominence of that syllable within a sentence (e.g. Hayes 1995; Liberman & Prince 1977; Prince 1983; Selkirk 1984). Let us take an example sentence "(Yes, I saw) five bright highlights in the sky tonight", in which all the target vowels are /aɪ/. We can posit a metrical structure shown in (1) (the judgment is the first author's). The more grid marks a syllable has, the more prominent that syllable is within a sentence.

(1) English sentential metrical structure: an example

Utterance			*			
PPhrase	*		*		*	
PrWd	*	*	*		*	*
foot	*	*	*	*	*	*
syllable	*	*	*	*	*	*
	five	bright	high	lights	(in the) sky	(to)night
# of marks	4	3	5	2	4	3

Several studies examined the relationship between jaw displacement patterns and sentential prominence patterns, and found a significant correlation between them (Erickson 2004a,b, 2010a,b; Erickson et al. 2012, 2014a, 2014b). These studies also found that larger jaw opening corresponds to higher F1 for low vowels (cf. Stevens 1998), which would arguably serve as the acoustic cue to listeners for large jaw opening. These observations are shown in Figure 3 for four speakers of American English producing the utterance "(Yes, I saw) five bright highlights in the sky tonight" (reduced syllables "in the" and "to" are not shown).
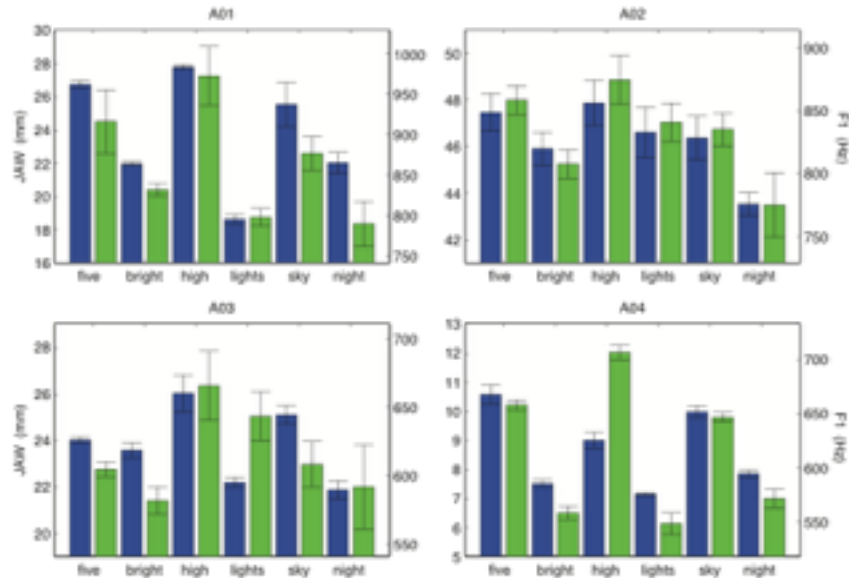
Figure 3. Jaw displacement (blue dark bars) and F1 measurements (green pale bars) for four American English speakers. The target sentence is "(Yes, I saw) five bright highlights (in the) sky (to)night." Based on Erickson et al. (2012), reproduced with permission.

In Figure 3, dark blue bars represent the degree of jaw displacement. The degree of jaw movement correlates very well with the number of metrical grids of each syllable shown in (1) (i.e. 4-3-5-2-4-3). The only exception is Speaker A04, who shows the biggest jaw opening on "five" instead of "high"—it is possible that this speaker assigns the Utterance-level grid mark on "five" instead of "high". The green, pale bars in Figure 3 show F1 of each syllable, which shows very close correlation with the magnitude of jaw opening ($r = 0.74$ for A01, $r = 0.5$ for A02, $r = 0.24$ for A03, $r = 0.70$ for A04). In summary, jaw displacement patterns reflect relative sentential prominence levels in English, and F1 may plausibly be their acoustic cues.
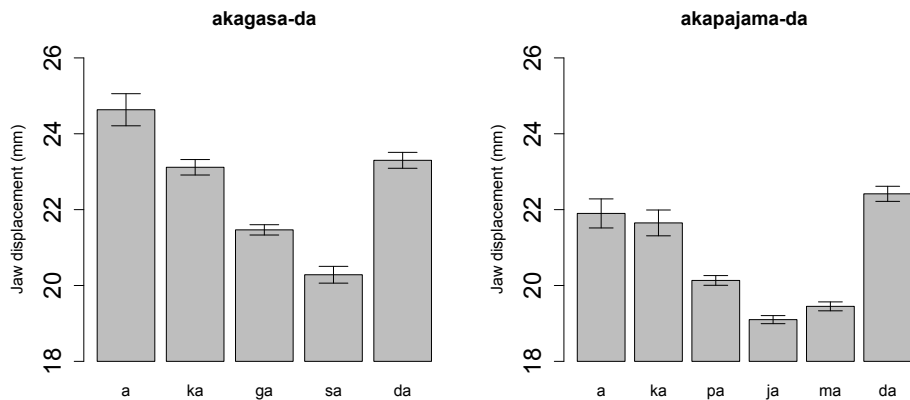
3.2. Dealing with vowel height effects

In the example sentence used in section 3.1, all the vowels were identical (=[aɪ]). The experiment controlled the vowel quality, because we know that vowel height affects degree of jaw opening as well (e.g. Keating et al. 1994). Then what would happen to "usual" sentences with different vowel qualities? Would we still observe that the sentential prominence pattern is reflected in the degrees of jaw opening? Our hypothesis is vowel height differences can *hide* the relationship between metrical prominence and jaw opening, but we can nevertheless *wash away*—or *normalize*—these effects by subtracting each vowel's specific jaw displacement factor.

In order to assess the metrical structure of an utterance which contains different vowels, a vowel normalization method was developed (Menezes & Erickson 2013; Williams et al. 2013). To simplify a bit, this algorithm calculates average jaw displacement values for each vowel, and subtracts each of these average values from raw values. Given this simple normalization method, it is possible to assess the metrical structure of an utterance,

independent of vowel quality.[5]

To illustrate, Figure 4 shows the sentence pair "Pat met Kip"/"Kip met Pat", with the raw data on top and normalized data on the bottom. The raw data for "Kip" has the smallest jaw opening whether it is in the initial or final position, because high (=closed) vowels show small jaw opening. Similarly, "Pat" always shows the largest jaw opening, because it contains an low (=open) vowel. However, once the vowels' effects are washed away using the algorithm, the two sentences have the same metrical pattern—with sentential stress on the final word (bottom two graphs). The metrical structure of the normalized jaw displacement

values for both sentences look similar to that proposed by Hayes (1995): the final element receives metrical prominence, which is indeed realized with large jaw opening.
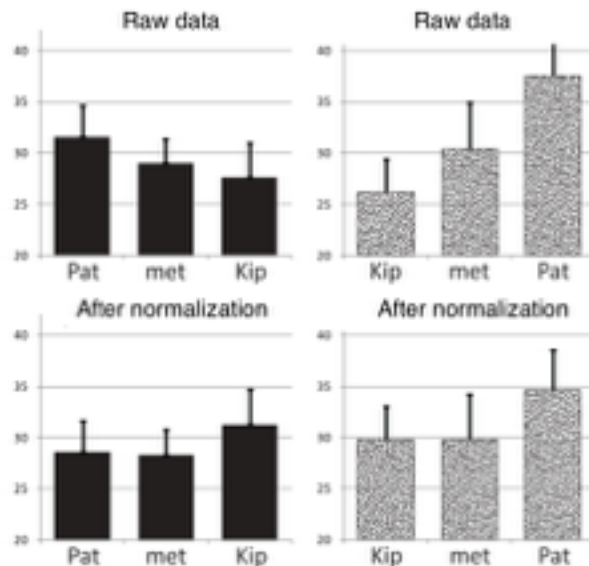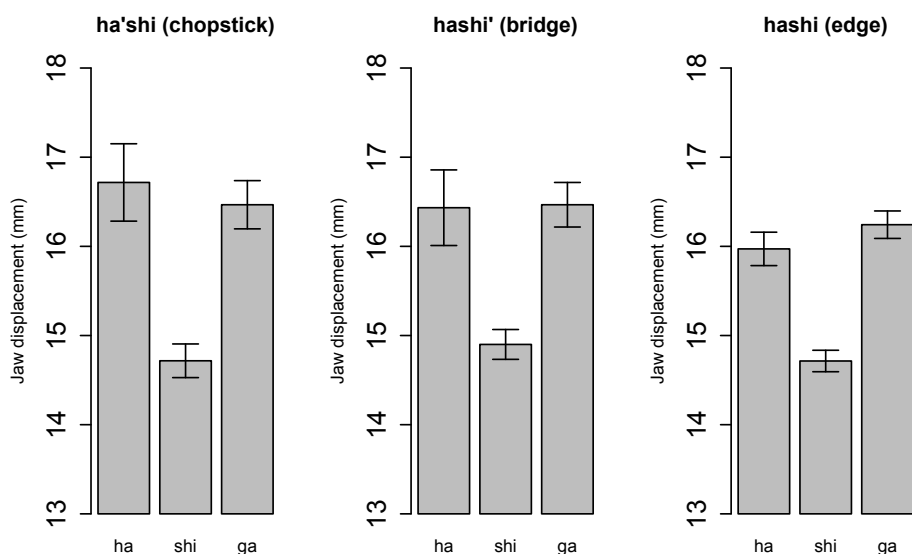

Figure 4. Top panel shows the raw jaw movement of "Pat met Kip" (left) and "Kip met

_____

[5] This algorithm assumes that the effects of vowel height and the effects of metrical prominence are independent; i.e. all kinds of vowels are equally affected by prosodic factors. Whether this assumption holds true is a topic of on-going research.

ha'shi (chopstick)   **hashi' (bridge)**   **hashi (edge)**

Pat" (right), and bottom panel shows the normalized data for the two sentences. Sentential stress is on the last word. Taken from Erickson et al. (2014c).

Furthermore, Erickson & Menezes (2013) show that the same algorithm can be used for similar three-word English sentences with initial sentential stress. These sentences with various vowel types also do not show consistent jaw displacement patterns, similar to the top two figures in Figure 4. However, once vowel effects are factored out using the normalization algorithm, we observe that jaw opening is consistently largest initially, which receives sentential stress.
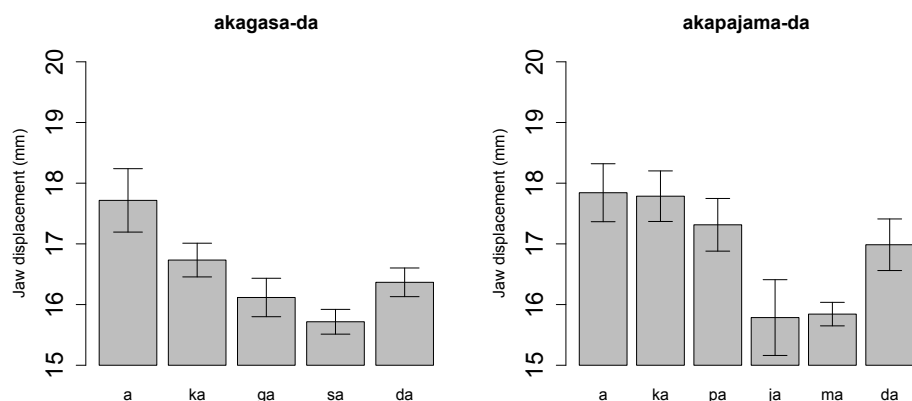
4. Jaw displacement patterns in Japanese
We now turn to some results on Japanese, which show that Japanese may have initial and final stress—apart from the well-known accentuation—within each phrase. As with the case of the English sentence discussed in section 3.1, we studied jaw movement patterns of Japanese sentences with the same vowel, /a/. Japanese is often considered as a pitch-accent language which lacks stress (e.g. Beckman 1986; Kawahara 2015), so would we expect jaw movement to be flat (or random)?

Figure 5, adopted from Kawahara et al. (2014), shows that (i) Japanese does have patterns of jaw displacement (i.e. it is neither flat nor random) and (ii) there seems to be large jaw opening at initial and final syllables.[6] Kawahara et al. (2015) moreover show that such syllables with large jaw opening exhibit higher F1 than other syllables, which can serve as cues for listeners to large jaw opening, just as in English.

Figure 5. Jaw displacement patterns of Japanese sentences consisting of syllables with [a] (two speakers). Taken from Kawahara et al. (2014). Reprinted with permission.

---

[6] See Kawahara et al. (2014) for longer sentences that involve multiple phrases, which show initial and final large jaw opening at each phrase.

Figure 6. Jaw displacement patterns of Japanese triplets that differ in accentual placement. Taken from Kawahara et al. (2014). Reprinted with permission.

Kawahara et al. (2014) also studied the effects of pitch accent on jaw movement patterns, using the famous triplet: /ha'si-ga/ 'chopstick-NOM' vs. /hasi'-ga/ 'bridge-NOM' vs. /hasi-ga/ 'edge-NOM'. These triplets show highly similar jaw displacement patterns, even though the pitch accent changes, as in Figure 6.

Why would Japanese show jaw opening at phrase-edge syllables, but not on accented syllables? There is a good reason to consider the Japanese edge syllables to be metrically strong.[7] In Tokyo Japanese, syllables at phrasal edges attract phrasal tones (Pierrehumbert & Beckman 1988), and in this particular sense, they can be considered metrically prominent (Jun 2014). Syllables with lexical accent are prominent in the theory of Jun (2014) and in other conceptions of Japanese metrical phonology as well, but there are phrases or sentences with only unaccented words. In such phrases or sentences, accented syllables are not prominent, simply because they do not exist. In addition, young Japanese speakers delete lexical accent in some particular contexts productively, producing unaccented words (Kawahara 2015). Therefore, it is only edge syllables that are *always* prominent, when we look at Japanese as a whole system.[8] It is not too mysterious, therefore, that Japanese phonology involves some abstraction in such a way that those syllables that are always prominent—to reiterate, syllables at the edges—receive articulatory prominence.

---

[7] We are very grateful to Sun-Ah Jun for extensive discussion on this point.

[8] As an anonymous reviewer reminded us, there is evidence from prosodic morphology that the first two elements within a word are phonologically strong in that they usually survive morphophonological truncation patterns (Ito & Mester 2015). See also Shaw (2007) for possible evidence for the metrical strength of foot-initial syllables in Japanese. These observations offer indirect evidence that even a stress-less language like Japanese shows metrical prominence in initial position. See also Bennett (2012) for a review of foot-initial prominence in many other languages.

5. Addressing some concerns

At this point, we address some alternative analyses of what we have shown so far, raised by an anonymous reviewer. First, to the extent that stressed syllables are longer (e.g. Fry 1955), the correlation between sentential stress and jaw displacement may be an epiphenomenon; the more stressed a syllable is, the longer it is, and the more time the speaker has to open their jaw. To address this question, we analyzed the correlation between jaw displacement and duration in both English and Japanese. For the English data shown in Figure 3, the acoustic data from the first three speakers were available, and the correlation coefficients between jaw displacement and duration were $r = 0.10$, $r = 0.38$, $r = 0.05$. Only the second speaker showed a significant positive correlation at the p-value of 0.04, which would not be significant after Bonferroni correction.[9]

Moreover, recall from Table 1 that emphasized [i] involves larger jaw displacement than unemphasized [i]. It is unlikely that the smaller jaw displacement of unemphasized [i] can be attributed to undershoot due to short duration, because for [i], the jaw does not have to travel long at any rate; i.e. undershoot of the jaw opening is unlikely to occur for [i].

An even more interesting pattern emerges in the Japanese data. Correlation between duration and jaw displacement for the two sentences in Figure 5 are shown in Figure 7 for three speakers. The acoustic data are based on Kawahara et al. (2015); the correlations are $r = -0.27$ ($p = 0.04$) for Sp S; $r = -0.21$ ($p = 0.1$) for Sp H; $r = -0.14$ ($p = 0.3$) for Sp Y.
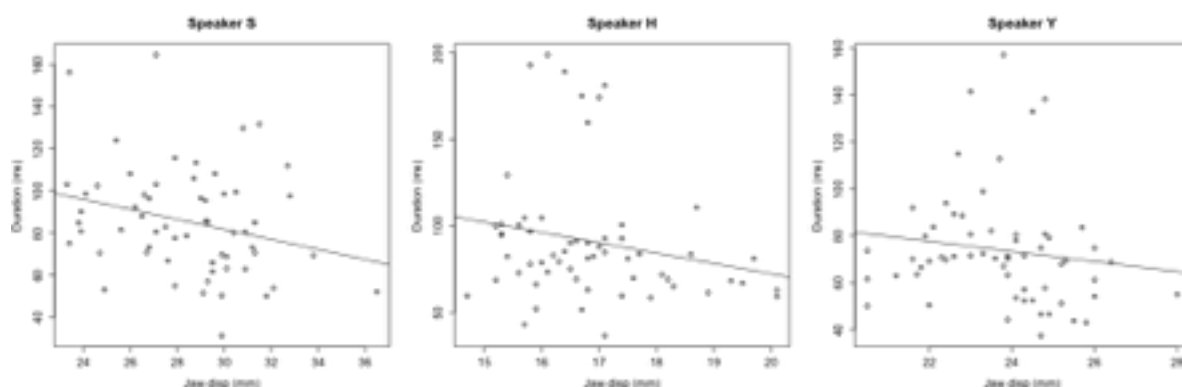

Figure 7: Correlation between duration and jaw displacement in Japanese.

We thus observe a weak *negative* correlation for all the three speakers. We conclude that it is hard to attribute large jaw opening to longer duration of the syllable under question.

An anonymous reviewer also raised the following question: can large jaw opening be explained away with recourse to domain-edge lengthening, which is arguably a universal reflex of motor control? In particular, many studies have shown that segments at phrase-edges are lengthened, which are often accompanied with more extreme gestures (Byrd & Saltzman 2003; Cho et al. 2006; Fougeron & Keating 1997 and many others). For Japanese in

---

[9] Is the correlation weak because the vowels in the stimulus sentences are diphthongs? What if we measure the duration of [a] excluding [ɪ]? Realistically speaking, it is difficult to measure the duration of one vowel within a diphthong, given the spectral continuity of the two vowels. However, it does point to an important follow-up experiment that should be conducted—these results should be replicated with a wider range of sentences with different vowels.

11

which phrase-edge syllables show large jaw opening, can this be because phrase-edge syllables are lengthened? This question indeed is a legitimate concern for Japanese, in which we observe large jaw opening at phrase-edges, but not for English in which the most prominent syllable is located at a phrase-medial position (*high* in Figure 3). There is an independent observation in English that large jaw displacement is not observed in final positions: "the final closing gesture [is] generally longer and slower but not more displaced" (Edwards & Beckman 1991: 369). Thus let us focus our discussion on Japanese.

Even in Japanese, jaw displacement patterns are, at least partly, independent of domain-edge lengthening. Recall from Figure 7 that large jaw opening cannot be attributed to lengthening of syllable durations in general. Further evidence comes from the observation that in Japanese, initial syllables are acoustically neither long nor strong. For illustration, Figure 8 shows the duration and intensity of each syllable for the target sentences shown in Figure 5.
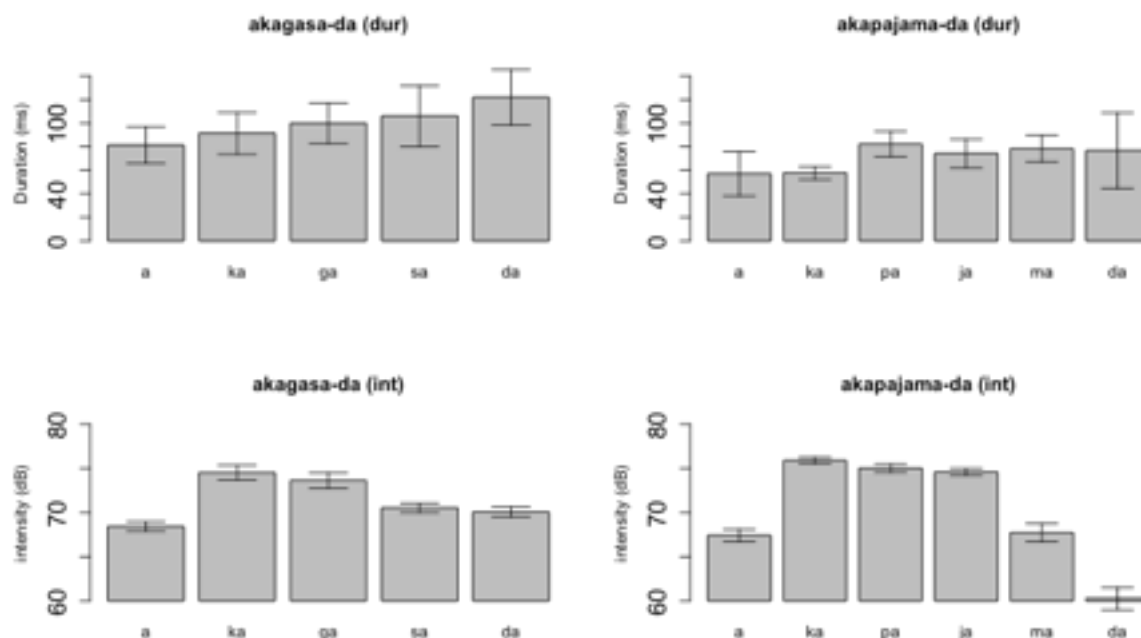

Figure 8. Duration and intensity of each vowels in the sentences shown in Figure 5. The data come from the acoustic analysis of Kawahara et al. (2015).

We observe that initial syllables in Japanese are neither long nor strong. Final syllables may be long due to utterance-final lengthening (top-left), but they show very low intensity (bottom-right, in particular), perhaps due to their heavy creakiness (Kawahara & Shinya 2008). We suspect that acoustically speaking, Japanese edge syllables are not very strong.

Ultimately, by looking at Japanese alone, it may be difficult to completely tease apart our hypothesis—the jaw displacement patterns reflect metrical prominence—from the alternative —that large jaw movement at phrase edges come from a universal domain-edge articulatory

lengthening effect.[10] The latter hypothesis, however, predicts that in all languages, large jaw displacement occurs at phrase-edges, both initially and finally, to the extent that phrase-edge strengthening is a universal mechanism based on motor control (Barnes 2006).

We suspect that this prediction does not hold in English, but it is important to keep looking at other languages. To this end, we now turn to preliminary observations about other languages.

6. Other languages

As we declared in the introduction, what we are illustrating in this paper is a new research program, and we hope to extend our studies beyond English and Japanese. Here we briefly mention our studies of other languages. Ongoing work with jaw displacement patterns in languages such as Spanish (Erickson et al. 2015), French (Erickson & Smith in preparation) and Mandarin Chinese (Iwata et al. 2015) show that these languages also have language-specific patterns of jaw displacement: French and Mandarin Chinese seem to have phrase-final large jaw opening, whereas Spanish may have phrase-initial large jaw opening. The observation about Spanish accords well with the observation that Spanish speakers systematically assign secondary stress on initial syllables (e.g. Hualde 2010).

Let us briefly illustrate the case of French, although we are yet to collect more data to make a more quantitative claim. A preliminary study shows that French has final prominence; jaw opening becomes larger over the course of a sentence, as shown in Figure 9. This observation accords well with the classification of French a being head/edge prominence (Jun 2014; Jun & Fougeron 2002).[11] In this model, the head is an Accentual Phrase final full vowel, the head acting as an edge marker. Thus, the French data is exactly what is expected, if metrical structure determines jaw movement patterns.
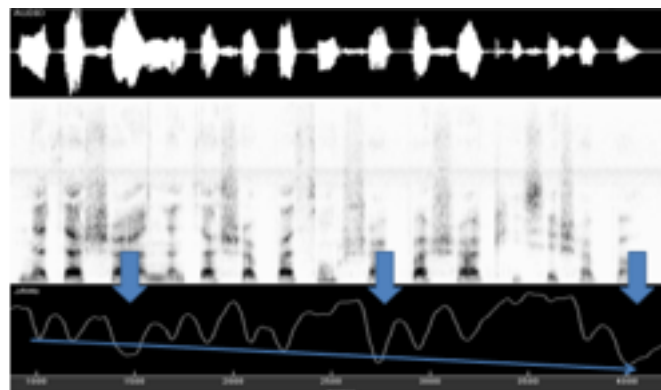

Figure 9. French jaw displacement patterns for one speaker of French. The sentence is "Natasha n'attacha pas son chat pacha qui s'echappa" (Natasha didn't tie her cat, Pasha, who

---

[10] Although in the latter hypothesis, we would have to posit that strengthened articulation can result in weak acoustic consequences, as in Japanese. See Barnes (2006) for related discussion.

[11] An anonymous reviewer pointed out that despite the head-finality of French, Fougeron & Keating (1997) found evidence for domain-initial lengthening in French. This is good evidence showing that domain-edge lengthening and large jaw opening are at least partly independent of one another.

escaped from her).

As an anonymous reviewer pointed out, there are two ways to implement this gradual decline: one is to posit that every prosodic level in French is right-headed, so that French has a metrical structure like (2):

(2)   Utterance   *
       Phrase   *   *
       Word   *   *   *
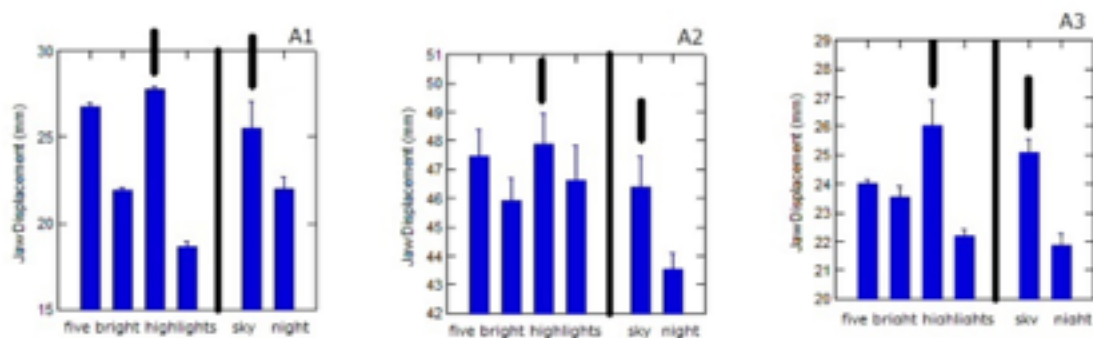   Syllable   *   *   *   *   *   *

Alternatively, we could just posit that final syllables are prominent, and posit a kind of gradient interpolation—or some sort of declination toward the beginning. See Fujimura (2003) for an idea for a declination function related to the second view.

The current proposal predicts that in no languages should the jaw movement pattern be flat or random (assuming that all languages have metrical structure); those syllables that are metrically prominent should show large jaw opening in every language. This is an empirically testable claim and needs to be examined in many languages. Combined with the discussion in section 5, one general lesson emerges: we need to collect jaw data to look at relationships between the jaw and metrical structure and between the jaw as well as other phonetic (articulatory or acoustic) parameters carefully, and this needs to be done in many languages.

7. Further implications

With what we have shown so far, we hope to have demonstrated that it is useful to study jaw displacement patterns as articulatory correlates of metrical structure. In addition to providing a new theoretical insight into how phonological metrical structure manifests itself phonetically, there are a few other benefits of this general research project. One is the study of second language acquisition: the jaw displacement patterns of the first language may well influence jaw displacement patterns when speaking a second language.

Japanese speakers of English tend to produce English sentences with large jaw opening sentence-finally, even in cases where there is no final stress (Erickson et al. 2014a; Wilson et al. 2012). Let us again take the sentence "(Yes, I saw) five bright highlights in the sky tonight". Recall that English speakers assign a strong-weak pattern to the final phrase "sky tonight". Japanese speakers, on the other hand, show large jaw opening at the final word. This comparison is shown in Figure 10.
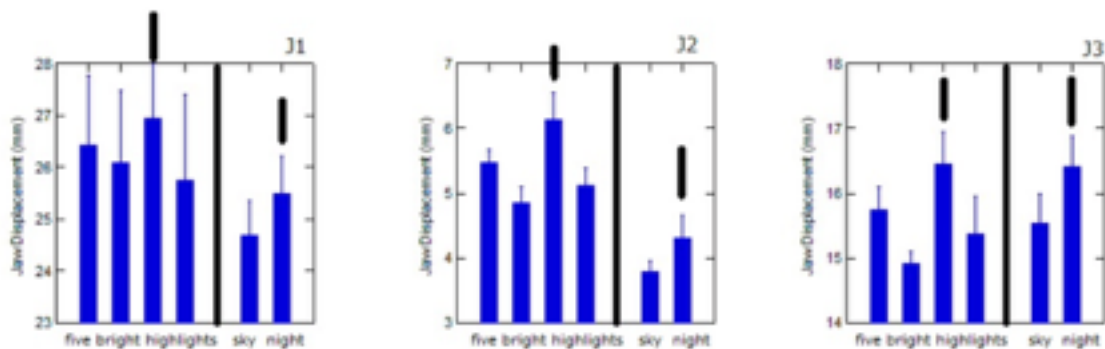
Figure 10: Jaw displacement pattern of English speakers (top) and Japanese speakers' L2 speech (bottom). The strongest jaw displacement in the first and second phrases are shown with thick black bars. Taken from Erickson et al. (2014a). Reprinted with permission.

The patterns in Figure 10 can be understood as phonetic L1 transfer in L2 speech; recall from Figure 5 that Japanese speakers show large jaw opening phrase-finally in their L1 speech, and it is natural to consider what we observe in Figure 10 to be a transfer effect from their native language. Wilson et al. (2012) show moreover that this transfer effect diminishes as L2 proficiency goes up. One of our ultimate goals is thus to use jaw movement as one of the measurements of L2 proficiency.

In addition, recent work by Wilhelms-Tricarico (2015) suggests that speech synthesis may be improved by first starting with the syllable as the basic unit, not consonant or vowel segmental units. Although still much detail needs to be worked out, understanding of the metrical nature of jaw movement can lead to better controls of parameters related to prosody, resulting in a number of achievements, including improved speech synthesis. See Wilhelms-Tricarico (2015) for a proposal for new speech synthesis design incorporating the research of the sort reported in this paper.

8. Concluding discussion
The new hypothesis that we are putting forward is very simple: in addition to properties produced at the larynx, metrical prominence manifests itself in the patterns of jaw movement. We think that this is an understudied hypothesis, and is worth more extensive exploration.

One question that often gets asked about this hypothesis is the following: when we are talking about metrical structure in terms of jaw movement, are we talking about the same metrical structure that governs phonological processes (like tonal alignment and stress assignment) and phonetic realizations, or are we talking about something slightly different? In answer to this question, we would like to start with the strong hypothesis that there is indeed one metrical structure which governs everything: phonology, acoustics, and jaw movement (see Bennett 2012 for related discussion and proposal). In English we seem to observe a very tight correlation between metrical stress and jaw opening, implying the isomorphism.

In Japanese, on the other hand, we did not find the effect of pitch accent, which is demonstrably assigned by a metrical foot (Kawahara 2015 and references cited therein). However, we argued that pitch accent has no effects on jaw movement, because pitch accent is not always present, whereas phrasal edge tones are always present; i.e. Japanese is an edge-

prominent language (Jun 2014). It may turn out in the end that we need to posit different metrical structures for, say, tonal distributions and jaw movement, just like other cases of "metrical inconsistency" in which tones and stress show evidence for conflicting metrical structuring (Vaysman 2009; cf. Bennett 2012).

To the extent that there is only one metrical structure, our proposal makes a testable prediction. To generalize, our proposal boils down to the thesis that there are multiple ways in which speakers express metrical structure, one of which is the jaw. This proposal thus predicts that when speakers are prevented from using their jaw—say, having a bite block or chewing gum—then, they would resort to some other articulator to express the metrical prominence instead. This prediction would be an interesting test for the general thesis pursued in this project.

To summarize, we proposed in this paper that it is worth investigating jaw as an articulator of metrical structure. We have presented evidence from English and Japanese that shows that jaw displacement patterns are neither random nor flat, but instead correlate well with metrical prominence. To the extent that jaw displacement patterns are articulatory correlates of metrical structure, we may be able to turn around and investigate the nature of the metrical organization of a particular language by studying jaw displacement patterns.

References
Barnes, J. 2006. *Strength and weakness at the interface: positional neutralization in phonetics and phonology*. Berlin: Mouton de Gruyter.
Beckman, M. E. 1986. *Stress and Non-Stress Accent*. Dordrecht: Foris
Beckman, M. E. & J. Edwards 1994. Articulatory evidence for differentiating stress categories. In P. Keating (ed.), *Papers in Laboratory Phonology* III, 7-33. Cambridge: Cambridge University Press.
Bennett, R. 2012. *Foot-conditioned phonotactics and prosodic constituency*. Doctoral dissertation, University of California, Santa Cruz.
Byrd, D., and E. Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31: 149–180.
Byrd, D., J. Krivokapic & S. Lee. 2006. How far, how long: On the temporal scope of phrase boundary effects. *Journal of the Acoustical Society of America* 120: 1589-1599.
Cho, T. 2006. Manifestations of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In L. M. Goldstein, D. H. Whalen & C. T. Best (eds.). *Laboratory Phonology 8*, 519-548. Berlin: Mouton de Gruyter.
Cho, T., J. McQueen & E. Cox. 2006. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35: 210-243.
de Jong, K. 1995. The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491–504.
Eady, S. J., W. Cooper, G. V. Klouda, P. R. Mueller, & D. W. Lotts. 1986. Acoustical characteristics of sentential focus: Narrow vs. broad focus and single vs. dual focus environments. *Language and Speech* 29: 233-251.
Edwards, J., M. E. Beckman & J. Fletcher 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89: 369-382.
Erickson, D. 1998. Effects of contrastive emphasis on jaw opening. *Phonetica* 55: 147-169.
Erickson, D. 2002. Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59: 134-149.

Erickson, D. 2003a. The jaw as a prominence articulator in American English. *Acoustical Society of Japan Fall Meeting* 311-312.

Erickson, D. 2003b Some effects of prosody on articulation in American English. In T. Honma, M. Okazaki, T. Tabata, S. Tanaka, *A New Century of Phonology and Phonological Theory, A Festschrift for Prof. Haraguchi*, 473-491. Tokyo: Kaitakusha.

Erickson, D. 2004a. On phrasal organization and jaw opening. *Proceedings From Sound to Sense, In Honor of Ken Stevens June 13, MIT.* 24, CD-ROM publication.

Erickson, D. 2004b. Metrical structure and jaw opening. *The 18th International Congress on Acoustics, Kyoto* (SP02). #0049.

Erickson, D. 2010a. An articulatory account of rhythm, prominence and phrasal organization. *Proceedings of Speech Prosody, Chicago*. #2006.

Erickson, D. 2010b. More about jaw, rhythm and metrical structure. *Acoustical Society of Japan Fall Meeting.* 103.

Erickson, D. & O. Fujimura. 1992. Acoustic and articulatory correlates of contrastive emphasis in repeated corrections. *Proceedings of the International Conference on Spoken Language Processing.* 835-837.

Erickson, D. & O. Fujimura. 1996. Maximum jaw displacement in contrastive emphasis. *Proceedings of the International Conference of Spoken Language Processing.* 141-144.

Erickson, D., O. Fujimura & J. Dang. 1999. Articulatory and acoustic characteristics of emphasized and unemphasized vowels. *Journal of the Acoustic Society of America* 106: 2241 (A).

Erickson, D, M. Hashi, & K. Maekawa. 2000. Articulatory and acoustic correlates of prosodic contrasts: A comparative study of vowels in Japanese and English. *Journal of the Acoustic Society of Japan* 56: 265-266.

Erickson, D. & K. Honda. 1996. Jaw displacement and F0 in contrastive emphasis. *Journal of the Acoustic Society of America* 99: 2494 (A).

Erickson, D., S. Kawahara, Y. Shibuya, A. Suemitsu, & M. Tiede. 2014a. Comparison of jaw displacement patterns of Japanese and American speakers of English: A preliminary report. *Journal of Phonetic Society of Japan* 18: 88-94.

Erickson, D., S. Kawahara, J. C. Williams, J. Moore, A. Suemitsu, A., & Y. Shibuya. 2014b. Metrical structure and jaw displacement: An exploration. *Proceedings of Speech Prosody 2014*: 300-303.

Erickson, D., S. Kawahara, I. Wilson, C. Menezes, A. Suemitsu, Y. Shibuya, & J. Moore. 2014c. Jaw displacement patterns as articulatory correlates of metrical structure. Talk presented at Phonetic Building Blocks of Speech, in Honor of John Esling, Sept. 18-21.

Erickson, D., K. Lenzo & O. Fujimura. 1994. Manifestations of contrastive emphasis in jaw movement. *Journal of the Acoustic Society of America* 95: 2822 (A).

Erickson, D., A. Suemitsu, Y. Shibuya, & M. Tiede. 2012. Metrical structure and production of English rhythm. *Phonetica* 69: 180-190.

Erickson, D., & C. Smith. in preparation. Jaw displacement and French metrical structure.

Erickson, D., J. Villegas, I. Wilson, & Y. Iguro. 2015. Spanish articulatory rhythm. *Acoustical Society of Japan Fall Meeting*. 319-322.

Eriksson, Anders. 2009. A typology for word stress and speech rhythm based on acoustic and perceptual considerations. Research Project: http://flov.gu.se/english/research/projects/typology.

Fougeron, C., & P. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101: 3728–3740.

Fujimura, O. 2003. Stress and tone revisited: Skeletal vs. melodic and lexical vs. phrasal. In S. Kaji (ed.) *Cross-Linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies*, 221–236. Tokyo: Tokyo University of Foreign Studies.

Féry, C. 2013. Focus as prosodic alignment. *Natural Language and Linguistic Theory* 31: 683–734.

Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 1765–1768.

Harrington, J., J. Fletcher, M.E. Beckman. 2000. Manner and place conflicts in the articulation of accent in Australian English. In M. Broe, J. Pierrehumbert (eds.), *Papers in Laboratory Phonology V*, 40-51. Cambridge: Cambridge University Press.

Hayes, B., 1995. *Metrical Stress Theory*. Chicago: The University of Chicago Press.

Hualde, José. 2010. Secondary stress and stress clash in Spanish. In Marta Ortega-Llebaria (ed.), *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, 11–19. Somerville, MA: Cascadilla Press.

Ito, J. & A. Mester. 2015. Word formation and phonological processes. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*, 363–395. Berlin: Mouton de Gruyter.

Iwata, R., D. Erickson, Y. Shibuya, A. Suemitsu. 2015. Articulation of phrasal stress in Mandarin Chinese. *Acoustical Society of Japan Fall Meeting*, 207-210.

Jun, S.-A. 2005. Prosodic Typology. In S.-A. Jun (ed.) *Prosodic typology: The phonology of*

intonation and phrasing, 430-458. Oxford: Oxford University Press.

Jun, S.-A. 2014. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In S.-A. Jun (ed.) *Prosodic typology II: The phonology of intonation and phrasing*, 520-539 Oxford: Oxford University Press.

Jun, S.-A & C. Fougeron. 2002. The realizations of the Accentual Phrase in French intonation. *Probus* 14: 147-172.

Kawahara, S. 2015. The phonology of Japanese accent. In Haruo Kubozono (ed.) *The handbook of Japanese phonetics and phonology,* 445-492. Mouton: Berlin.

Kawahara, S., D. Erickson, J. Moore, A.Suemitsu, & Y. Shibuya. 2014. Jaw displacement and metrical structure in Japanese: The effect of pitch accent, foot structure, and phrasal stress. *Journal of Phonetic Society of Japan* 18: 77-87.

Kawahara, S., D. Erickson, & A. Suemitsu. 2015. Edge prominence and declination in Japanese jaw displacement patterns: A view from the C/D model. *Journal of Phonetic Society of Japan* 19: 33-43.

Kawahara, S. & T. Shinya. 2008. The intonation of gapping and coordination in Japanese: Evidence for Intonational Phrase and Utterance. *Phonetica* 65: 62-105.

Keating, P. A., B. Lindblom, J. Lubker & J. Kreiman. 1994. Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics* 22: 407–422.

Kent, R. D. & R. Netsell. 1971. Effects of stress contrasts on certain articulatory parameters. *Phonetica* 24: 23-44.

Ladd, D. R. 2008. *Intonational phonology, 2nd Edition*. Cambridge: Cambridge University Press.

Liberman, M. & A. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249-336.

Loevenbruck, H. 1999. An investigation of articulatory correlates of the accentual phrase in French. *Proceedings of 14th International Congress of Phonetic Science*: 667-670.

Macchi, M. 1985. *Segmental and suprasegmental features and lip and jaw articulations*. Doctoral dissertation, New York University.

Macchi, M. 1988. Labial articulation patterns associated with segmental features and syllable structure in English. *Phonetica* 45: 109-121.

Maekawa, K., S. Kiritani, H. Imagawa, & K. Honda. 1998. Effects of focus on mandible movement. *Journal of the Acoustical Society of America* 54: 259-260.

Menezes, C. 2003. *Rhythmic pattern of American English: An articulatory & acoustic study*. Doctoral dissertation, Ohio State University.

Menezes, C. 2004. Changes in phrasing in semi-spontaneous emotional speech: Articulatory evidences. *Journal of Phonetic Society of Japan* 8: 45-59.

Menezes, C. 2015. Comparing the metrical structure of a digit sequence in American English: Articulatory vs. acoustic analysis. *Journal of Phonetic Society of Japan* 19.

Menezes, C. & D. Erickson. 2013. Intrinsic variations in jaw deviation in English vowels. *Proceedings of International Congress of Acoustics, POMA* (19). #060253.

Menezes, C., B., Pardo, D. Erickson, and O. Fujimura. 2003. Changes in syllable magnitude and timing due to repeated correction. *Speech Communication* 40: 71-85.

Oshimat, K.& V. L. Gracco. 1992. Mandibular contributions to speech production. *Proceedings of International Conference on Spoken Language Processing, Banff.* i92_0775.

Perkell, J. 1969. *Physiology of speech production*. Cambridge: MIT Press.

Pierrehumbert, J. B. & M. E. Beckman. 1988. *Japanese Tone Structure*. Cambridge: MIT Press.

Plag, I., G. Kunter & M. Schramm 2011. Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics* 39: 362–374.

Poser, W. J. 1984. *The phonetics and phonology of tone and intonation in Japanese*. Doctoral dissertation, MIT.

Prince, A. 1983. Relating to the grid. *Linguistic Inquiry* 14: 19–100.

Selkirk, E. 1980a. Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff & M.Kean (eds.), *Juncture*, 107–129. Saratoga, CA: Anma Libri.

Selkirk, E. 1980b. The role of prosodic categories in English word stress. *Linguistic Inquiry* 11: 563–605.

Selkirk, E. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.

Selkirk, E., & J. Katz. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87: 771-816.

Shaw, J. 2007. /ti/~/tʃi/ contrast preservation in Japanese loans is parasitic on segmental to prosodic structure. *Proceedings of the 16th ICPhS*: 1365–1368.

Stevens, K. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Stone, M. 1981. Evidence for a rhythm pattern in speech production: Observations of jaw movement. *Journal of Phonetics* 9: 109-120.

Tilsen, S. 2009. Multitimescale dynamical interactions between speech rhythm and gesture. *Cognitive Science* 33: 839-879.

Summers, W. V. 1987. Effects of stress and final consonant voicing on vowel production: articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82: 847–863.

Vaysman, O. 2009. *Segmental alternations and metrical theory*. Doctoral dissertation, MIT.

Westbury, J. & O. Fujimura. 1989. An articulatory characterization of contrastive emphasis. *Journal of the Acoustical Society of America* 85: S98.

Wilhelms Tricarico, R. 2015. Text to speech synthesis using syllables as functional units of speech. *Journal of Phonetic Society of Jap*an 19.

Williams, J. C., D. Erickson, Y. Ozaki, A. Suemitsu, N. Minematsu, & O. Fujimura. 2013. Neutralizing differences in jaw displacement for English vowels, *Proceedings of International Congress of Acoustics, POMA* (19). #060268.

Wilson, I., D. Erickson, & N. Horiguchi. 2012. Articulating rhythm in L1 and L2 English: Focus on jaw and F0. *Acoustical Society of Japan Fall Meeting*, 319-322.